Keystone Genes in Evolving Genetic Networks

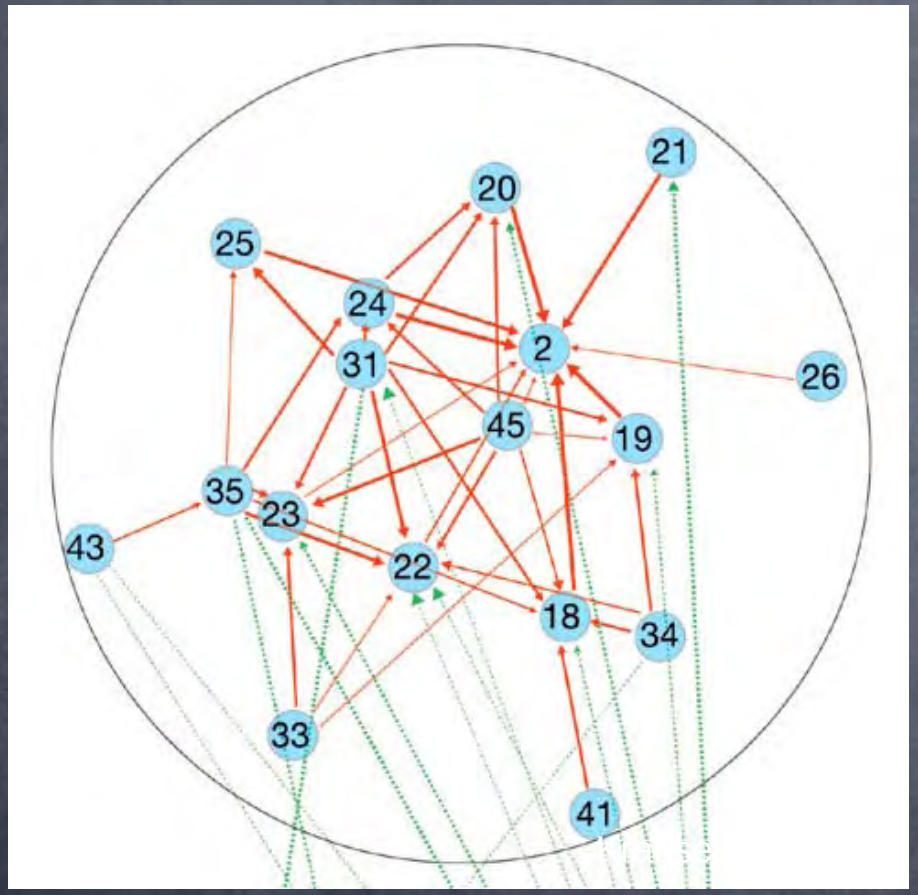
Stephen Proulx, Iowa State University

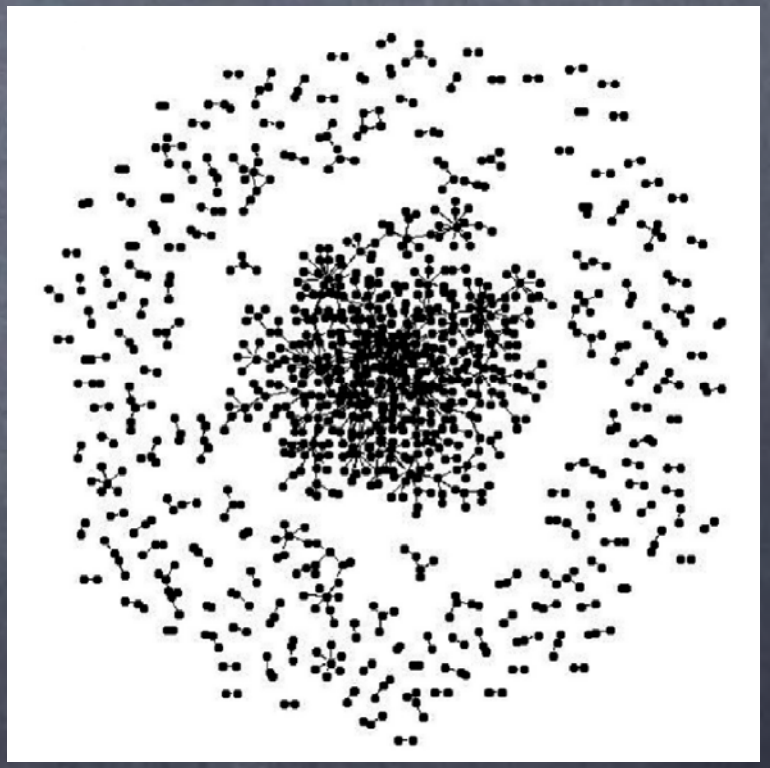
Collaborators

- Patrick Phillips, U Oregon
- Daniel Promislow, U Georgia
- Hong Hua, ISU
- Michael Smiley, ISU
- Bills: NIH and ISU

Biological Networks

Biological Networks

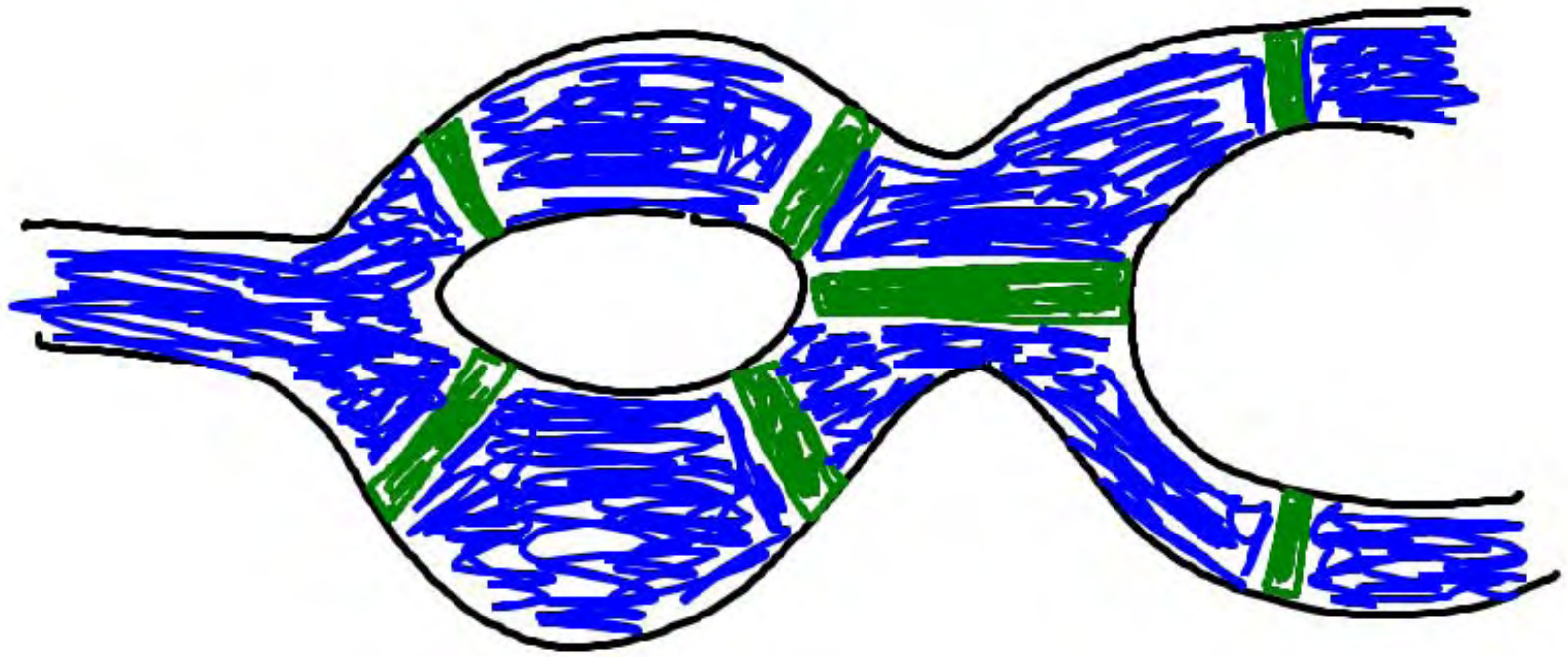




Why use networks?

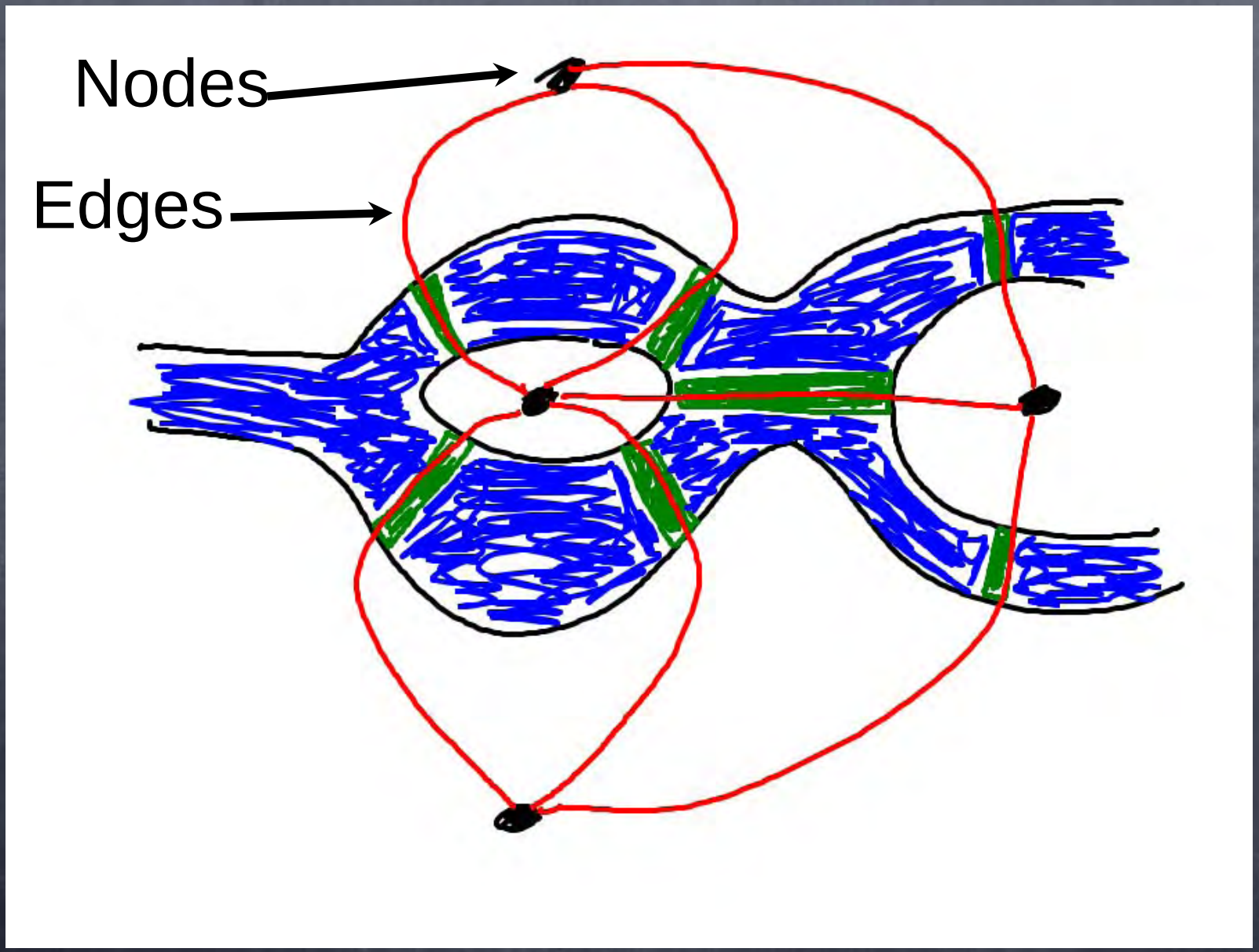
- Represent system by nodes and links
- Statistical methods for analysis exist
- Look for general results

Early Network Theory

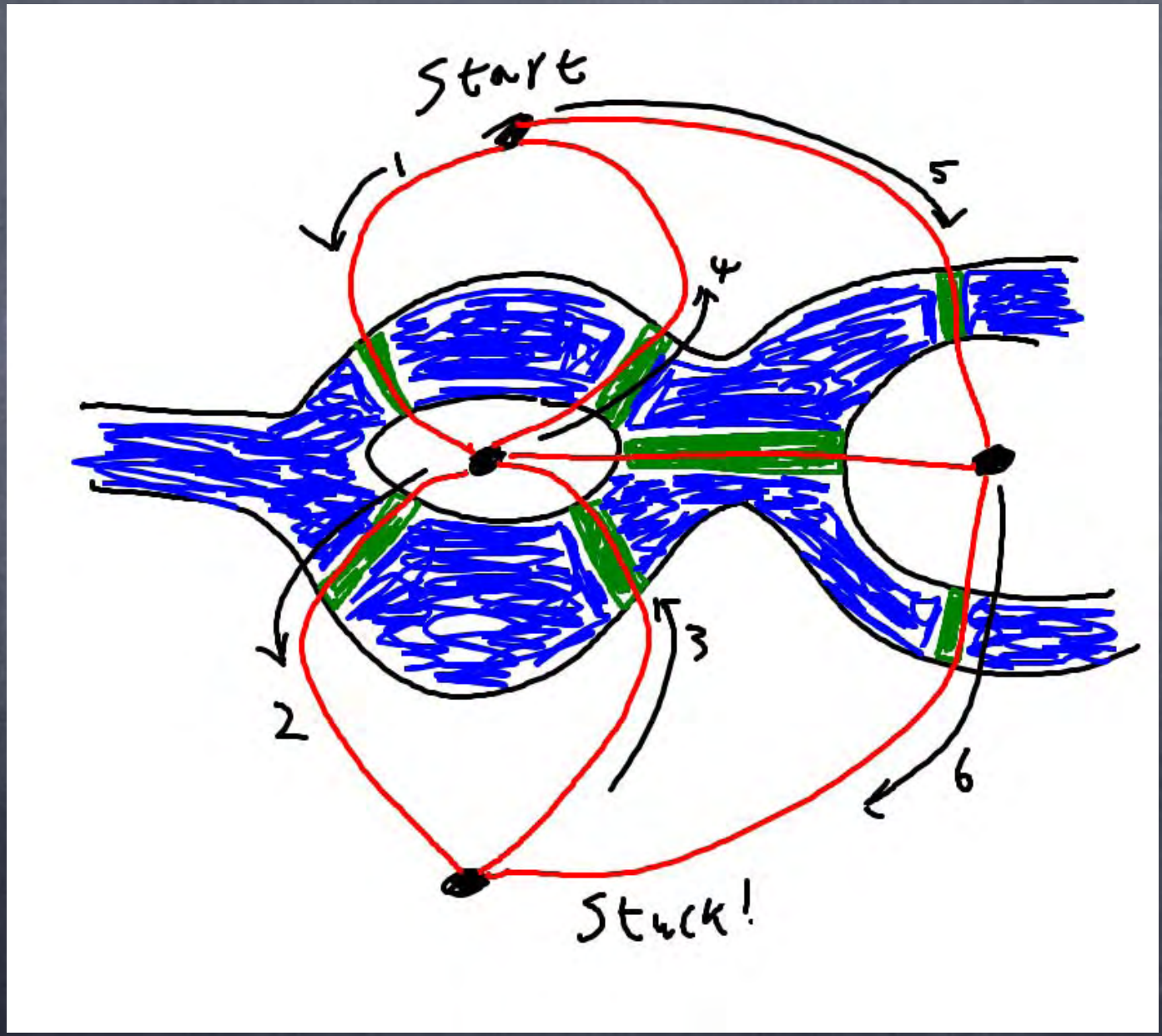


The bridges of Königsberg

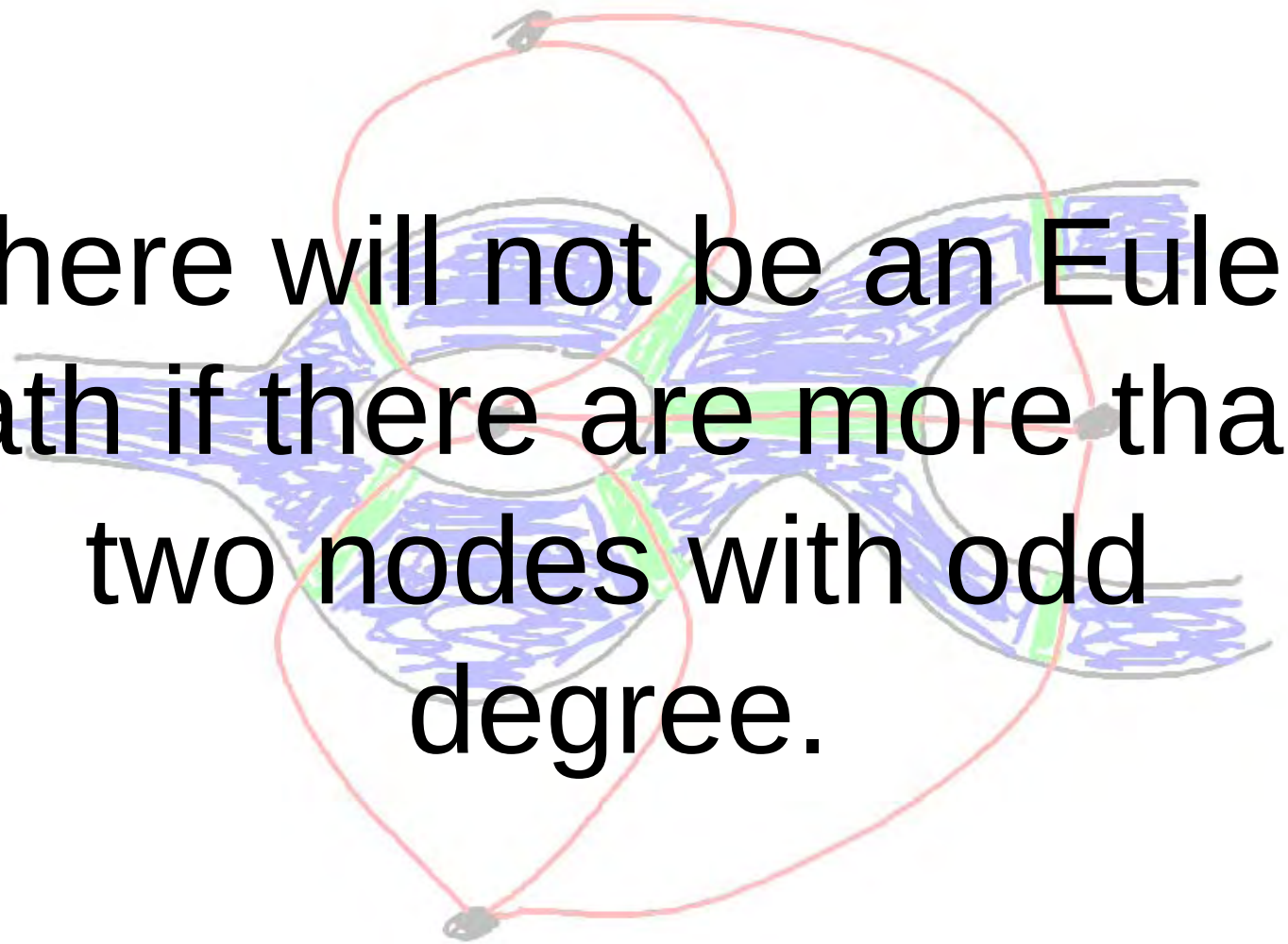
Paths Across Bridges



Possible Path

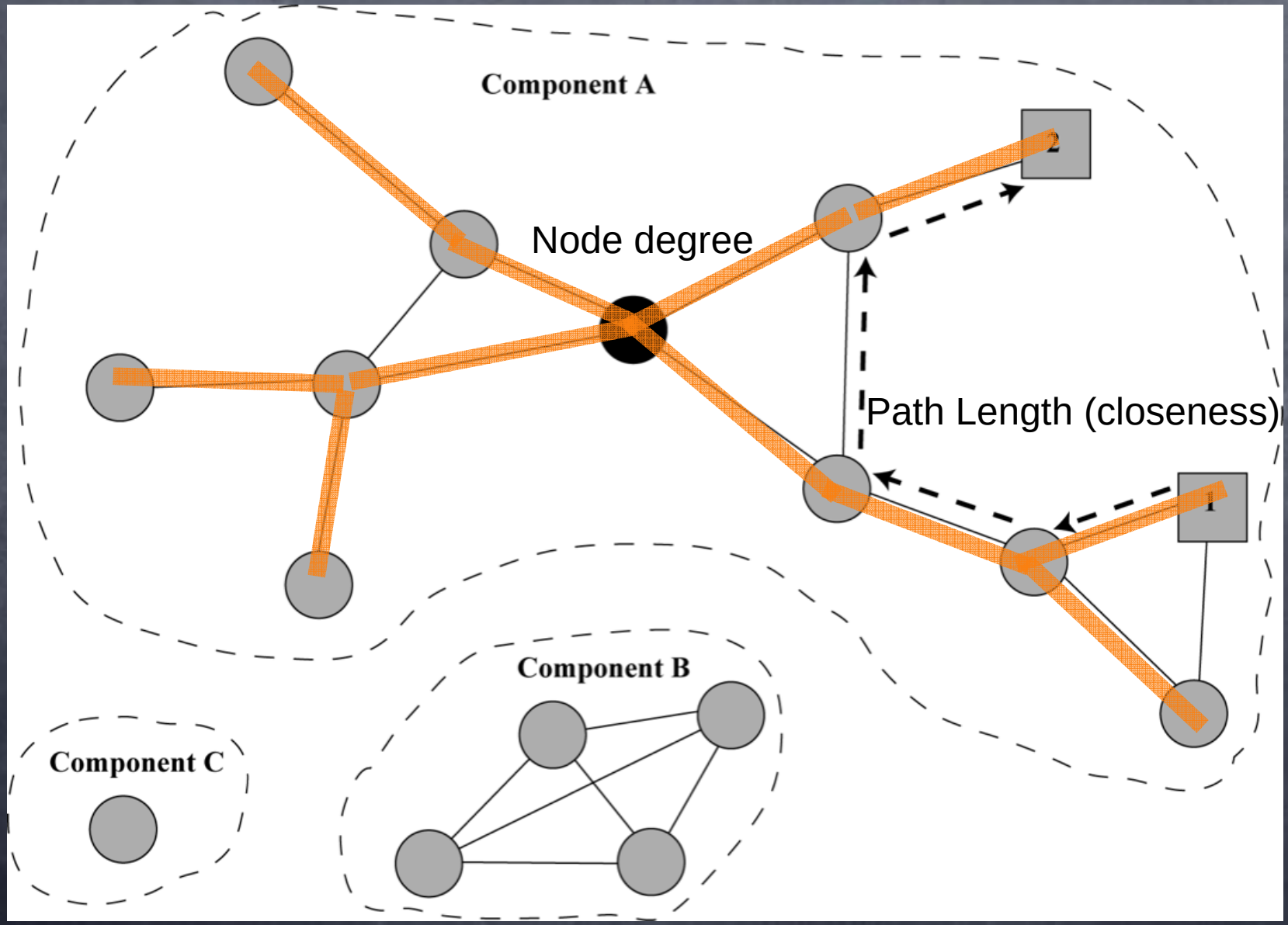


Euler's Result



There will not be an Euler path if there are more than two nodes with odd degree.

Network Position



Network Approach

Question: Can network continue to function after removal of a node?

Does position in the network alone tell us something about how network function is related to changes at that node?

ROBUSTNESS: Insensitivity of function when the system is perturbed.

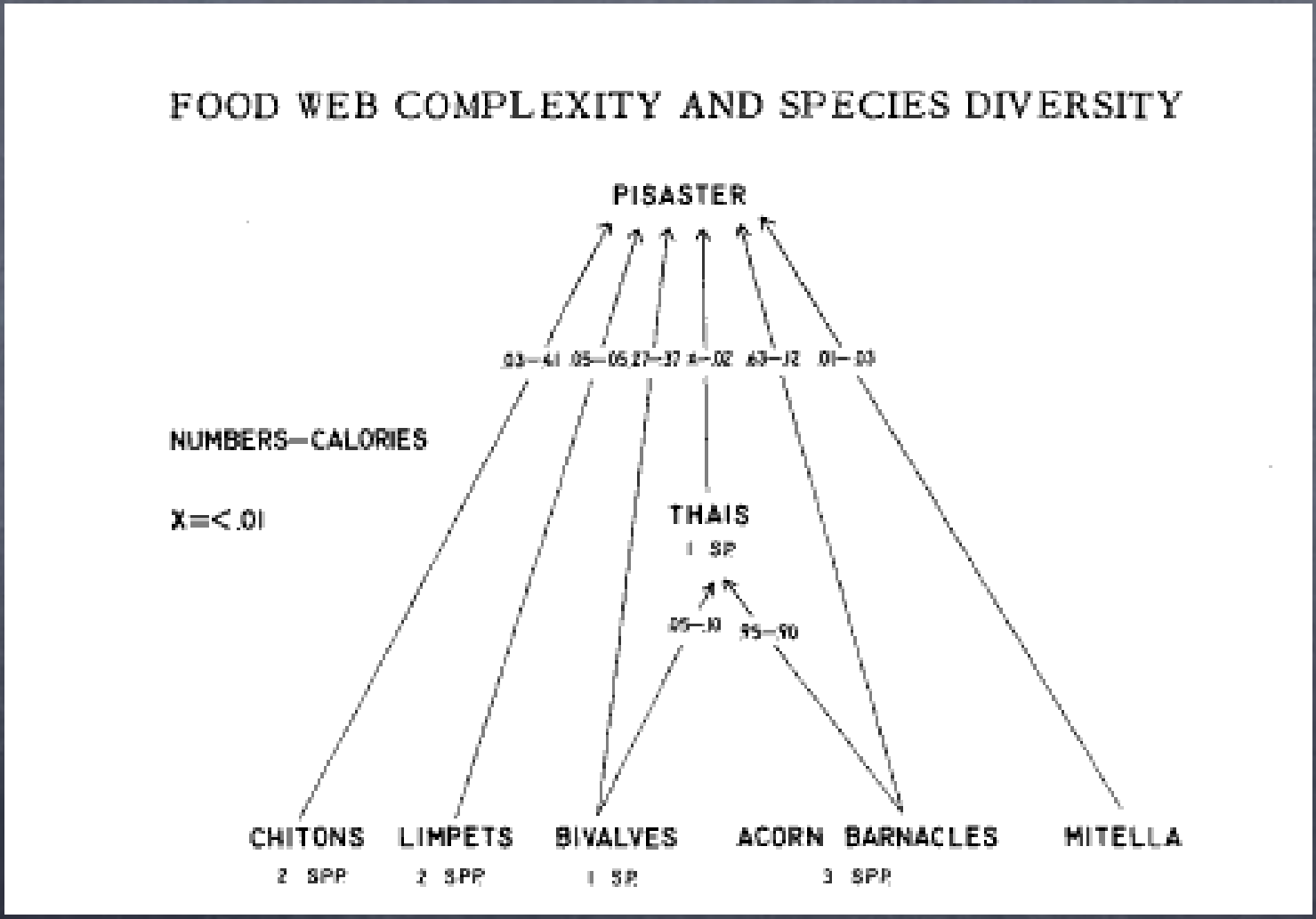
Network Keystones

- Keystone species: Species with large effect on community
- Typically has many interactions
- But, may owe importance to specific ecological interactions

Network Approach

- Start with complex system
- Determine players and links in network
- Calculate network statistics
- Use network statistics to make predictions about network function

Pisaster web



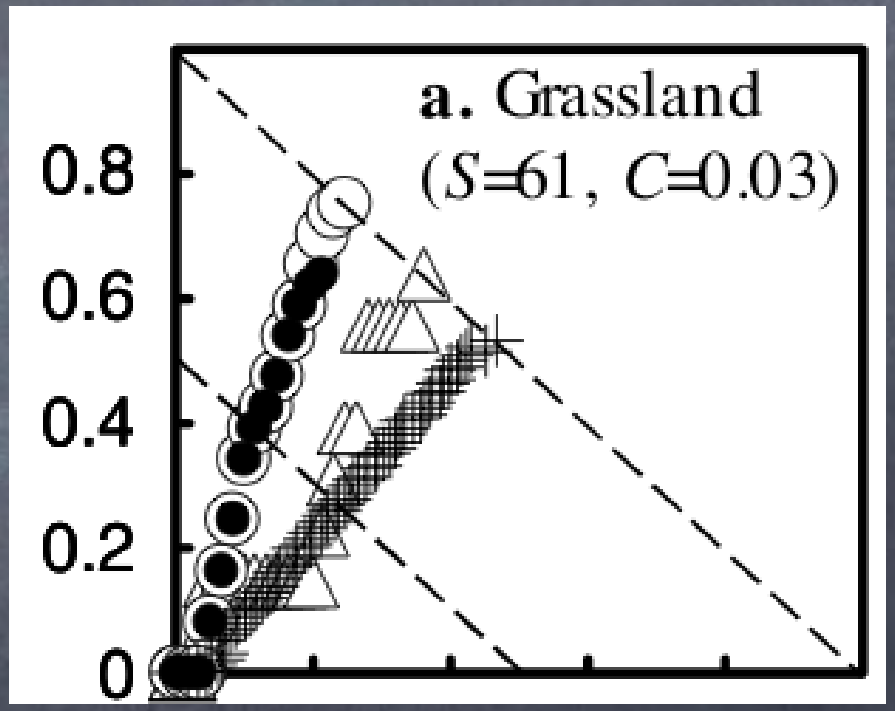
Paine, 1966

What is the right measure of topology?

- Number of species interactions?
- Known to be related to information flow
- Predicts species extinctions?

What is the right measure of topology?

- Try number of species interactions
- Known to be related to information flow
- Is correlated with number of secondary extinctions

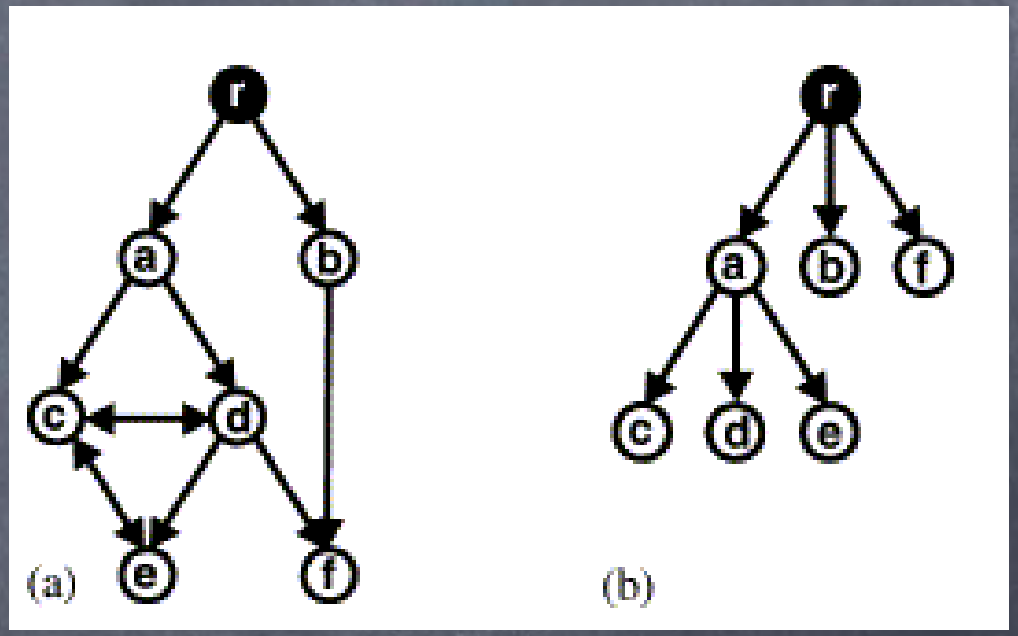


Species Removed

Dunne et al., 02

What is the right measure of topology?

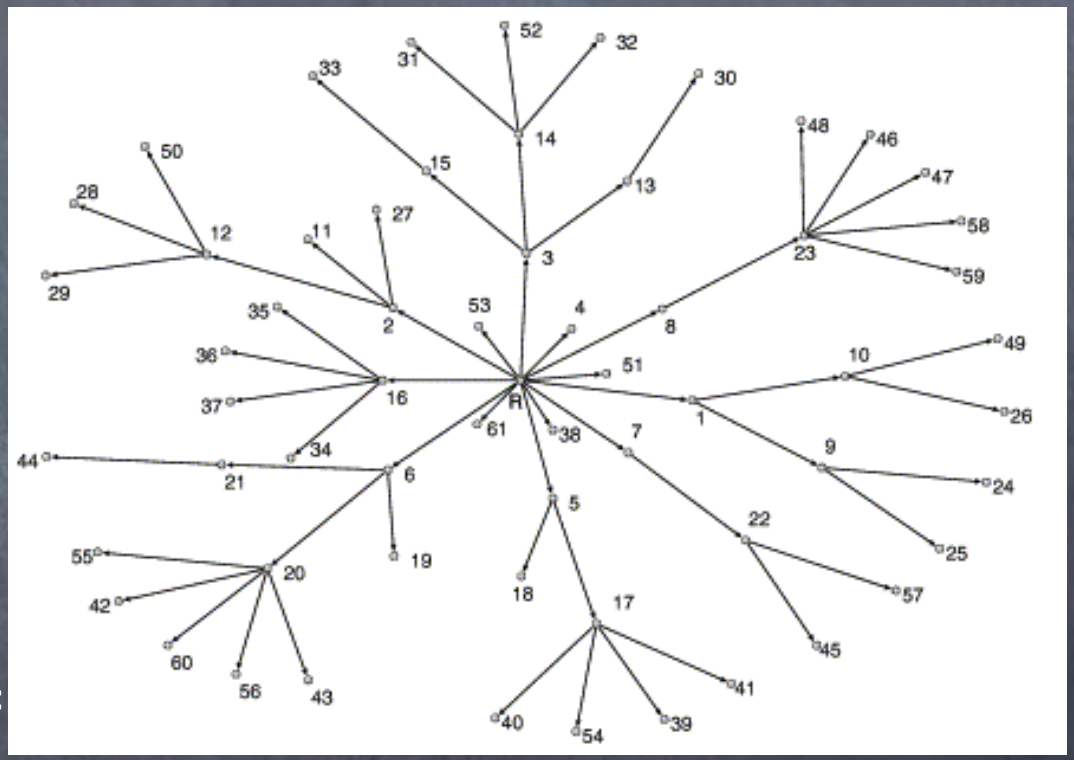
- Dominating tree
- Relates node to dependent species
- Can be correlated with degree



Allesina & Bodini, 04

What is the right measure of topology?

- Dominating tree
- Relates node to dependent species
- Correlated with degree
- Better prediction of extinction



Allesina & Bodini, 04

Ecosystem Robustness

- Ecosystem is more sensitive to loss of highly connected species
- Dominating nodes even more important
- However, some weakly connected species cause many extinctions

Ecosystem Robustness

Keystone species: Species that has larger effect on community composition than its biomass would suggest (Paine).

Keystone species, network version: Ecosystem is not robust to the loss of keystone species

Gene Networks

- Nodes are genes or gene products
- Interact through transcriptional control
- Protein interactions
- Metabolic interactions
- Larger scale interactions (e.g., cell-cell, tissue, organ, etc)

Evolution of Robustness

- Robustness— Also known as CANALIZATION or REDUNDANCY.
- Robustness is a result of the reduction of variance in a trait that is exposed to a perturbation.
- Genetic Robustness
- Environmental Robustness

Network Robustness

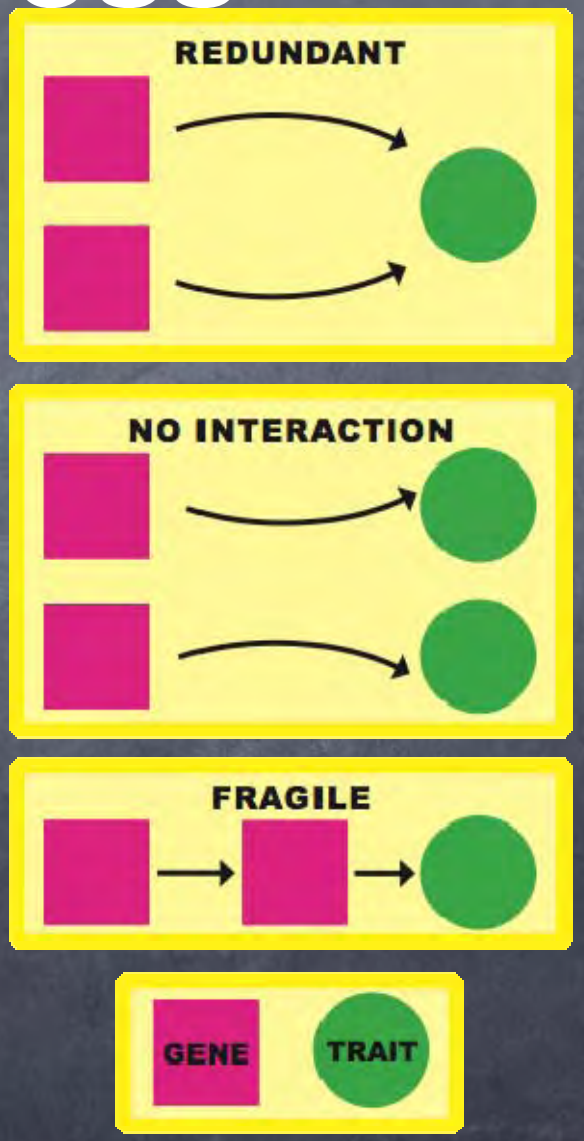
Strength of selection on robustness is related to the fitness load

$$\lambda \leq \frac{L}{1-L} \approx L$$

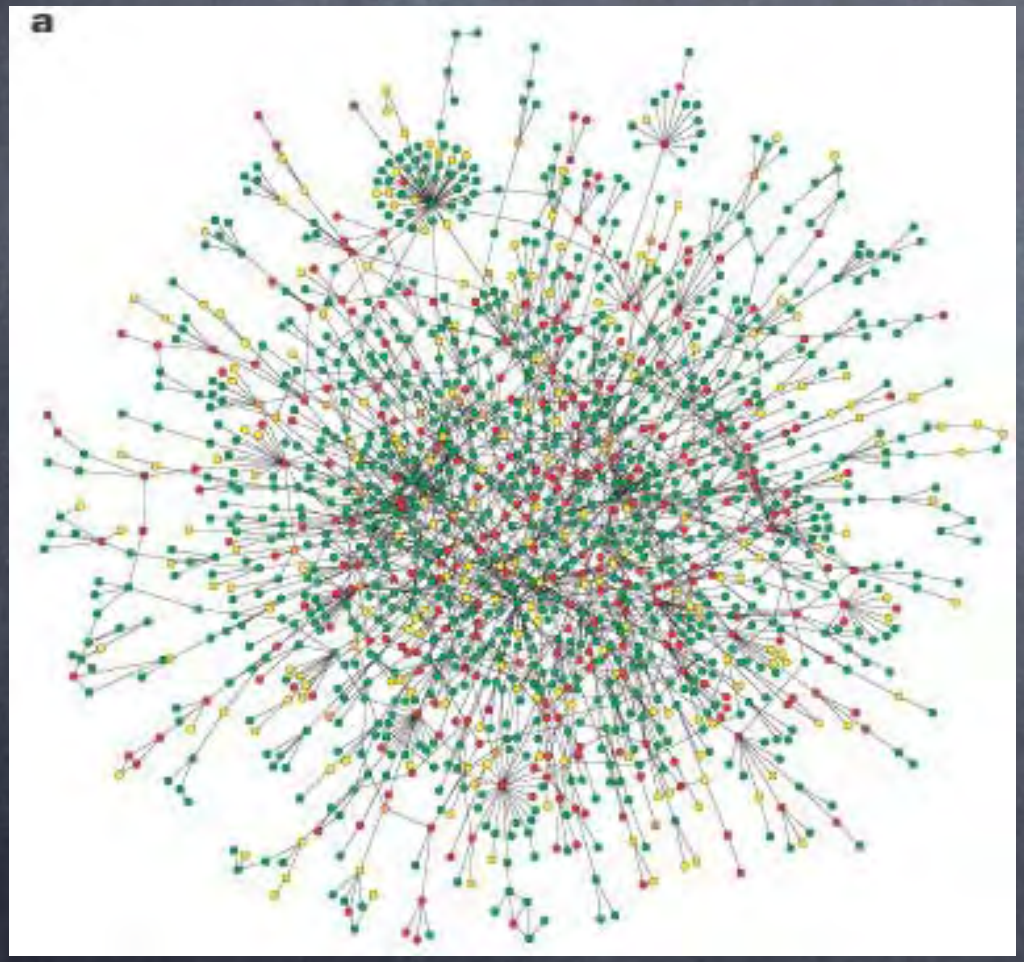
Proulx & Phillips 2005

Epistasis and Robustness

- Evolution of genetic robustness is related to fitness topology
- Genes that have more physical interactions have more potential to evolve genetic robustness



Protein Interaction Network



Jeong et al., 2001
Hahn & Kern 05

- Nodes are proteins
- Links are physical interactions
- Lethal genes are more connected and more central

Testing predictions with yeast expression data

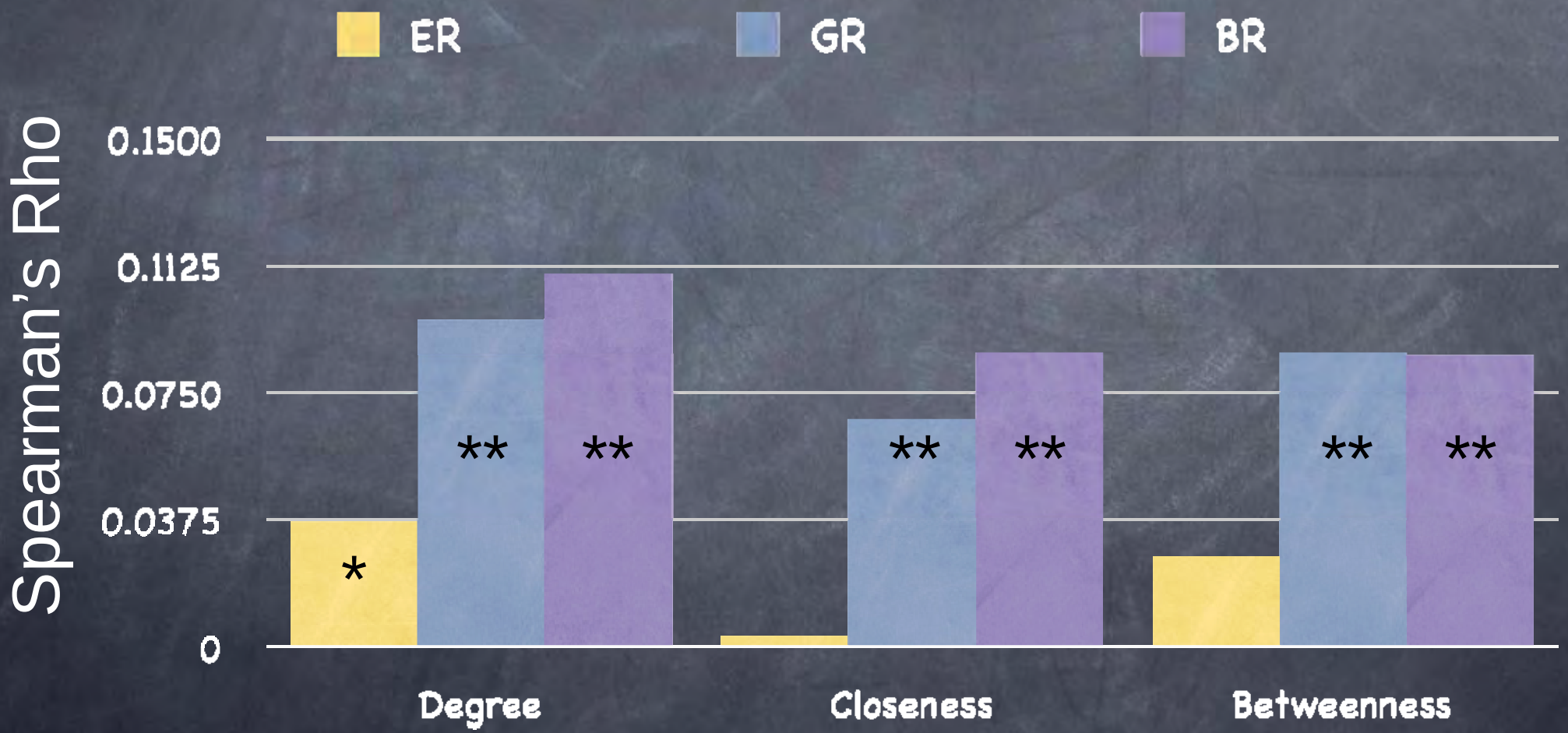
- Three data sets
  - Environmental shocks (167/35) (ER)
  - Gene knockouts (276) (GR)
  - Wild yeast strains (30) (BR)
- Measured expression using microarray of ~5000 genes

Gasch et al, 200; Hughes et al 2000, Nuzhdin unpublished data

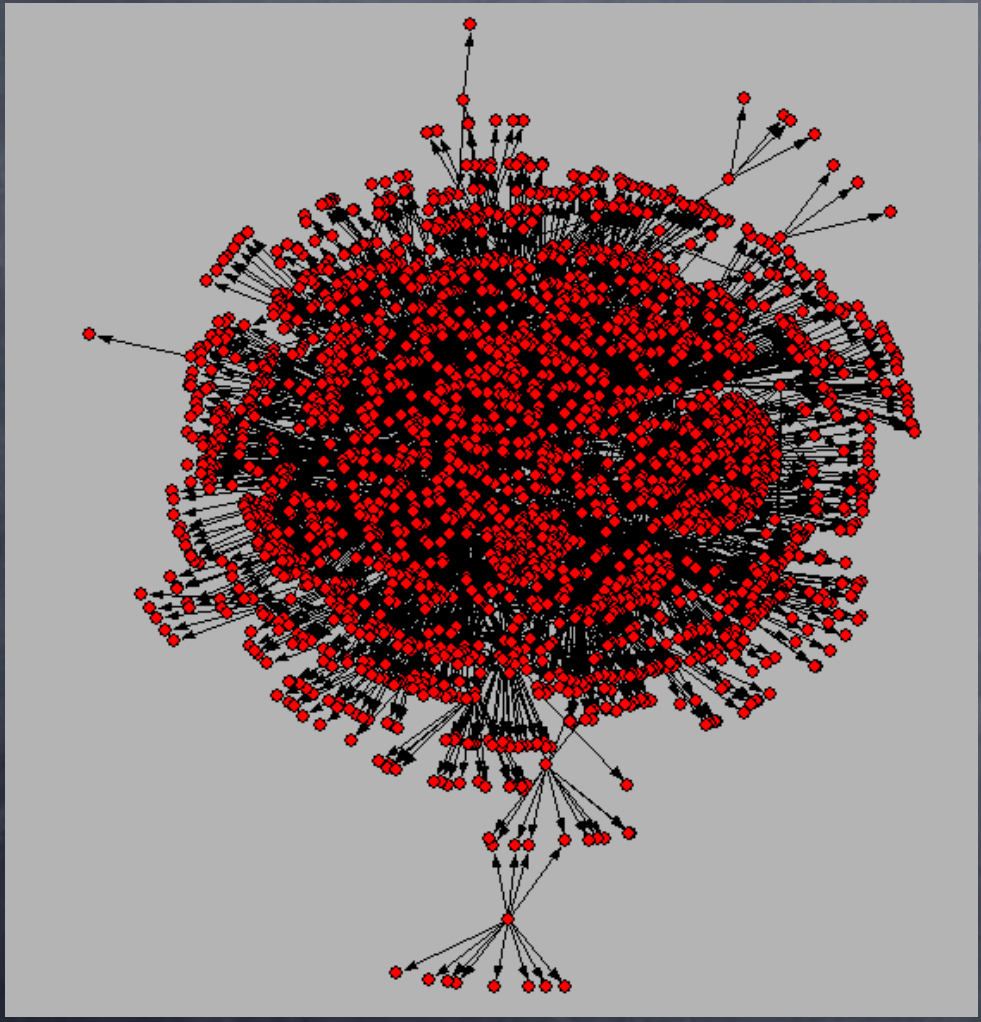
Three Robustness Measures

- ER– Environmental Robustness
- GR- Genetic Robustness (knockouts)
- BR- genetic Background Robustness

Network Position



Transcription Network



- Transcription factors
- Directional network
- In and Out degree

Transcription Degree

	ER	GR	BR
Kin	-0.093 **	-0.088 **	-0.050 *
Kout	NS	NS	NS

** p<10^-10

* p<0.0005

Transcription Degree

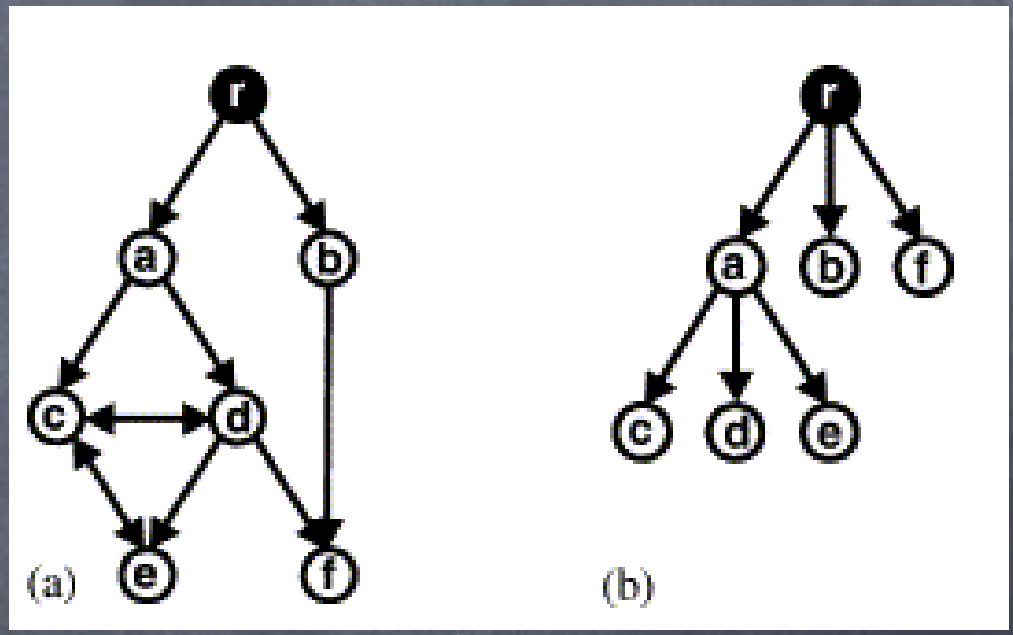
- Out-degree unrelated to robustness
- In-degree correlated with robustness
- Genes controlled by more transcription factors are less robust
- Protein connectivity significantly associated with robustness when in-degree is controlled for

Network Position

- High degree proteins have more robust expression
- High centrality is correlated with genetic robustness (not environmental)
- Genes with more binding sites are more plastic under any perturbation
- Can explain about 10% of variation in robustness. Significant but not predictable.

Finding Keystone Genes

- Can we use notions of energy flow and dominating nodes?
- Gene networks don't have a basal "resource"
- Can we use other aspects of topology like network motifs



Conclusions

- Robustness in ecosystems and gene networks is related to network position
- Central proteins buffer more genetic noise
- More transcription interactions means more plasticity
- Need to uncover additional important features of network position

Collaborators

- Patrick Phillips, U Oregon
- Daniel Promislow, U Georgia
- Hong Hua, ISU
- Michael Smiley, ISU
- Bills: NIH and ISU